Ph.D. Thesis Synopsis

Title

Identification Of Genetic Risk Factors Associated With Myocardial Infarction

February - 2023

Reg. No. MU/EXB/Ph.D./CR.96/A.Z./2017-2018/E.13, Dated: 29-07-2017

Registration Date : 25-01-2016

Submitted by

Nagarjuna Pasupuleti

Department of Applied Zoology

Mangalore University

Mangalore

Research Supervisor

Dr. Mohammed S Mustak

Professor

Department of Applied Zoology

Mangalore University

Mangalore

Place of Investigation

Department of Applied Zoology, Mangalore University, Mangalagangothri

Mangalore – 574199

Ph.D. Thesis Synopsis

Reg. No. MU/EXB/Ph.D./CR.96/A.Z./2017-2018/E.13

Research Scholar:
Nagarjuna Pasupuleti
Dept. of Applied Zoology
Mangalore University
Mangalore

Research Supervisor:
Dr. Mohammed S Mustak
Professor
Dept. of Applied Zoology
Mangalore University
Mangalore

Title: Identification Of Genetic Risk Factors Associated With Myocardial Infarction.

Objectives:

- To do GWAS in a sub-set of MI case and control samples using Illumina microarray platform
- To identify variants responsible for MI from GWAS data
- To identify variation in selected candidate genes, which may be responsible for susceptibility to develop MI
- Validation of MI associated variants from GWAS in a case-control cohort

1. INTRODUCTION

1.1 Research Background:

A heart attack which is the main reason behind heart tissue damage also scientifically known as myocardial infarction (MI). This kind of situation occurs due to the decrease in the blood flow or if the blood flow stops to a specific part of the heart. MI induces a most specific symptom which is chest pain or discomfort in the chest. This pain basically made the person uncomfortable and for which the patient started panicking. This particular pain is not only felt by the patient in the chest but it can move from chest to shoulder, arm, back, neck as well as sometimes in the jaw (Erdmann et al., 2010). Usually the pain occurs in the center part or sometimes left side of the chest and present there for more than a few minutes and then after a while it starts travelling towards the shoulder, arms and so on. After the attack the patient basically feels a heartburn like situation including that shortness of breath, feeling of fainting, nausea, tired feeling as well as cold sweat. In many previous studies it has been reported that women mostly feel neck pain, arm pain or may feel tired during this MI attack instead of any discomfort in chest or chest pain (Yamada et al., 2008). Studies have been previously reported regarding some clinical practice where it has been demonstrated that individuals of a certain age can be encountered with multiple TRF for MI without significant angiographic evidence. Presently the idea of pathophysiology of MI actively supports through modern clinical techniques to the notion. In the larger picture a heart attack basically occurs when one of the coronary arteries of the heart suddenly get blocked or due to slow blood flow face the shortage of blood. The typical reason behind this sudden blockage in the coronary artery is thrombus or the formation of a blood clot which in regards reduces the blood flow in the heart (Ivanova et al., 2017). These blood clots mainly forms inside the coronary artery which influence the blood flow as well as sometimes the coronary arteries get narrowed by fatty deposits (plaques) which build up next to the inside walls of blood vessels (atherosclerosis) make the passage more smaller and for which blood cannot flow properly in the arteries and disturbs the system of blood circulation (Wang et al., 2014). A person with low blood pressure is more prone to a MI as when the heart is beating it can be problematic to circulate the proper amount of blood in the heart as well as for low blood pressure, slow blood flow in the coronary artery can take place which in regards induces MI in a person. On the other hand it has been demonstrated and proved that the requirement of oxygen is greater than its supply then too a heart attack can take place and that too without formation of any blood clots (Kontou et al., 2018).

Atherosclerosis is a multi-decade pathological process involving complex interactions between genetics and environmental factors whereas the estimated heritability of CAD and MI is 50% to 60%. The heritability of each individual risk factor is indicated as a % in the figure 1. A study has been done by the genome-wide association where it has been identified successfully and revealed about a large number of factors that majorly contribute to the risk of myocardial infarction at the population level. The findings of the study has been reported based on a family where an autosomal dominant inheritance pattern for myocardial infarction has not been found. In these cases, a particular genetic factor worked as a inducer of MI which in regards elevates the risk of myocardial infarction so massively that nearly each and every individual possessing the risk-conferring gene, i.e., each of the child of an affected person, actually goes on to have a myocardial infarction. Even though it has been possible in these families, as in other perspective, to identify chromosome segments associated with a heritable risk of myocardial infarction, the genes and genetic variants responsible for the autosomal dominant inheritance of myocardial infarction in these families have not yet been disclosed.

Myocardial infarction (MI) indicates the death of the heart muscle tissue (myocardium) caused by ischaemia. This term is basically known as the lack of oxygen delivery to myocardial tissue (Schunkert et al., 2010). Depending on the result of ECG and also from some evidence it has been reported that MI can be classified into three different categories which includes unstable angina or Coronary Spasm, ST elevation myocardial infarction (STEMI) and Non-ST elevation myocardial infarction (NSTEMI). The detailed view about all of the tree MI is described below (Shiffman et al., 2005).

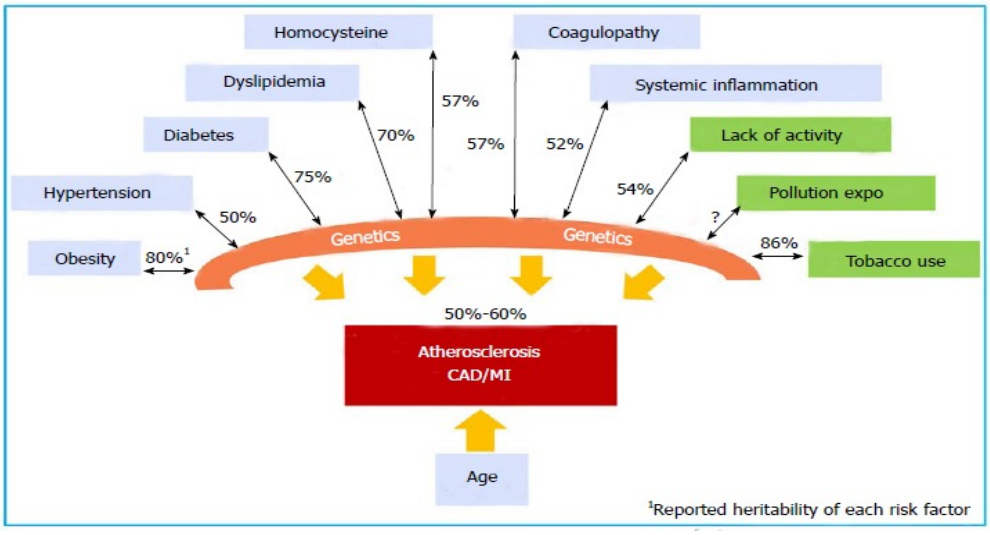

Figure 1: The genetic and environmental factors involved in MI

1.1.1 Unstable Angina or Coronary Spasm:

It has been reported that the formation of spasm of a coronary artery induces the to heart attack or myocardial infarction on a frequent basis. Specifically at the time of coronary spasm, the coronary arteries usually constrict or spasm the on and off which in regards cause the less amount of blood supply to the heart muscle which is known as ischemia and leads to heart attack. This particular type of MI can take place in people without significant coronary artery disease. Generally the longer the coronary artery spasm will be the more the chance of occurring heart attack. If it occurs for a long period of time that will definitely lead to heart attack or myocardial infarction (Matsuoka et al., 2015).

1.1.2 ST elevation myocardial infarction (STEMI):

A ST elevation myocardial infarction or STEMI mainly occurs when by the formation of blood clots the coronary artery gets completely blocked and for which a large portion of the heart muscle receives less amount of blood or stops receiving blood. This specific kind of MI is considered as the most serious heart attack which causes significant damage not only to the heart muscle but also can harm the person as well (Matsuoka et al., 2015; Winham et al., 2015) . Patients with STEMI may get familiar with a range of symptoms and experience crushing retrosternal or left sided chest pain or discomfort in the chest including to those typical symptoms to isolated dyspnoea, syncopal attacks, malaise and breathlessness or shortening in breath. Diabetic patients, persons with a elderly age and patients on NSAIDS are prone to suffer silent myocardial infarction or a minimal range of heart attack. It has been concluded in many previous studies that these patients are usually found to survive with cardiogenic shock, hypotension, arrhythmias and conduction blocks and acute left ventricular failure (Gore et al., 2015).

1.1.3 Non-ST segment elevation myocardial infarction (NSTEMI):

NSTEMI which basically stands for Non-ST-elevation myocardial infarction also well known as a non-STEMI. This is also a type of heart attack which is considered as a significant type of MI. NSTEMI is totally different from STEMI, the most common and usual type of heart attack which does not affect the person's heart that much and for that heart muscle gets less damaged. So, it can be clearly said that NSTEMI is less harmful than STEMI and also the electrocardiogram or ECG representation shows a distinct kind of waveform in case of NSTEMI which is dissimilar to STEMI (Gore et al., 2015).

For the management of patients with MI, many advanced procedures and approaches have been developed such as thrombolytic therapy and interventional therapy to relieve the patient from the severe pain. It is already known that MI has become a major problem worldwide, so it is very important to modify the diagnostic as well as the treatment methods for better outcome. People are facing MI issues from almost the last two decades and till now it has been reported that the main risk factor for the formation of MI or heart attack is genetic susceptibility and non-genetic factors which includes hypertension, obesity, diabetes and lifestyle as well. About the reason behind MI many studies have reported that smoking is the major reason to influence and induce these attacks and some previous study also showed that smoking is not only the major risk factor for the onset of MI but also it directs to a poor prognosis for MI (Manchanda et al., 2014). On the other hand, it has also been reported that a family history with premature coronary artery disease plays the key role in the development of MI which directly signifies a genetic correlation in between the sufferers.

As a myocardial infarction or MI positive family history known to be the most identical cardiovascular risk factor. So, to prevent the risk of MI in a family throughout, therefore the newly added European guidelines recommend intensified primary anticipation for the siblings and children of the persons who already have had an myocardial infarction. Previously the genetic relation between the family members were unknown and also the genes underlying the heritable component of MI were not familiar to us. Presently or may be from last some years the modification and development of innovative molecular genetic methods and especially the arrival of genome-wide association (GWA) studies in the field of diagnosis, has helped to direct to the detection of various genetic variants that are solely linked with an prominent risk factor of MI. Previous studies have demonstrated that the most strongest genetic effect which is specifically responsible for the risk of myocardial infarction is a significant segment on chromosome 9p21.3. In the present situation for myocardial infarction, this particular locus has been replicated almost in more than 45 000 cases and 85 000 controls. Though this locus or the region of the chromosome does not contain any specific gene which indicates any amino acid or particular protein. So, interestingly till now the underlying pathogenetic mechanism of this genetic hierarchy remains unclear. Based on the genetic point of view, it can be clearly mentioned that in most of the candidates the gene analyses have implicated among many molecular variants of an individual gene, only to one or a few. In the year of 2007, the breakthrough came in with the very first usage of genome-wide association (GWA) studies. The GWA type of analysis has the ability to characterize

approximately 2.4 million genetic markers in case of each matter (Ivanova et al., 2017). The significant systematic association of these markers in almost thousands of patients suffering with myocardial infarction and in healthy control matters as well, has facilitated the proper detection of gene segments for which the risk of atherosclerosis gets influenced and leads to MI beyond any doubt. Reports have been revealed that till date, there are the involvement of eleven genes or eleven chromosome segments which are designated as the risk of myocardial infarction.

Therefore the present study has designed with the following objectives

- To do GWAS in a sub-set of MI case and control samples using Illumina microarray platform
- To identify variants responsible for MI from GWAS data
- To identify variation in selected candidate genes, which may be responsible for susceptibility to develop MI
- Validation of MI associated variants from GWAS in a case-control cohort

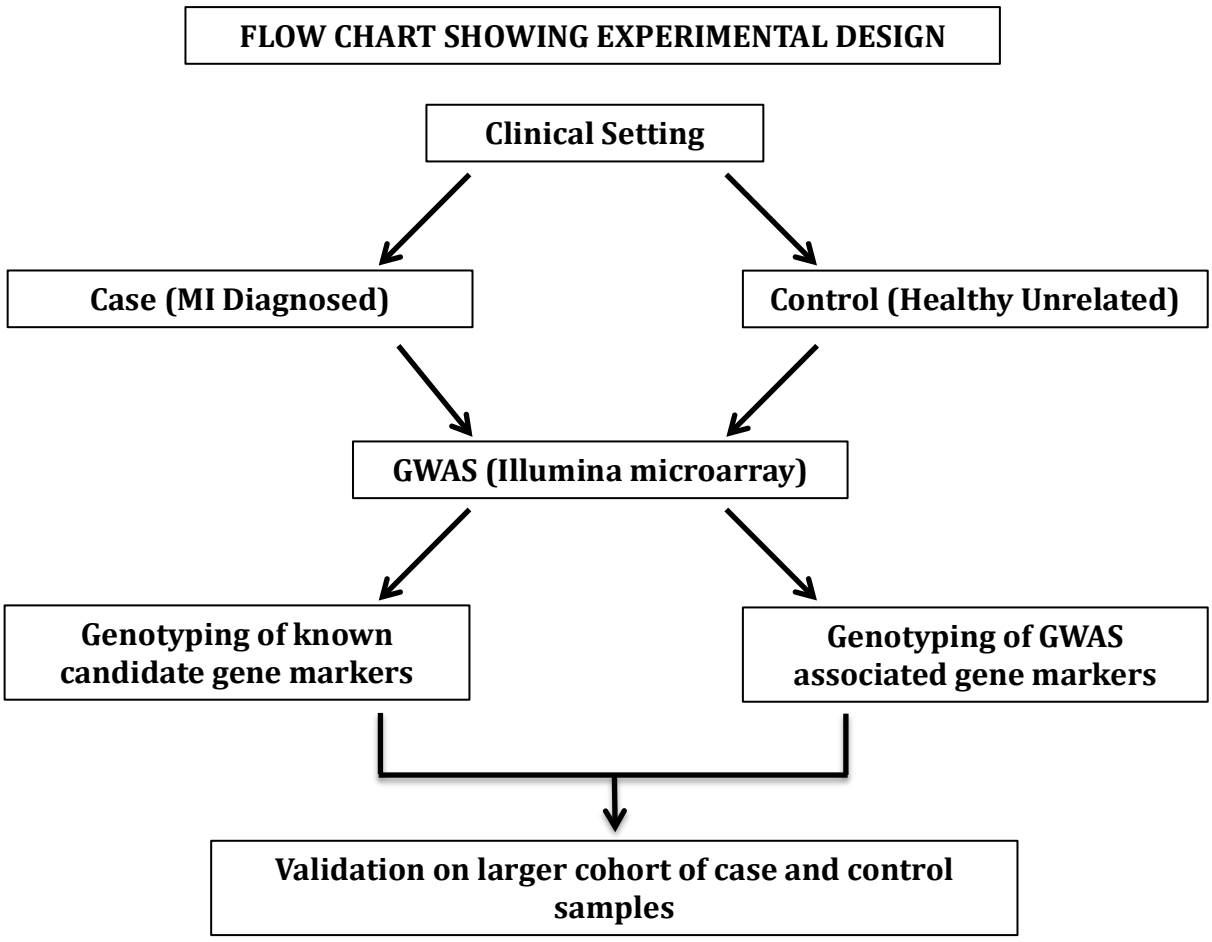
FLOW CHART SHOWING EXPERIMENTAL DESIGN
Clinical Setting
Case (MI Diagnosed)
Control (Healthy Unrelated)
GWAS (Illumina microarray)
Genotyping of known candidate gene markers
Genotyping of GWAS associated gene markers
Validation on larger cohort of case and control samples

Figure 2 : Flowchart showing the experimental design

2. MATERIALS AND METHODOLOGY :

2.1. Sample Collection: Clinically diagnosed myocardial infarction and control samples were collected from Dakshina Kannada, Karnataka with appropriate institutional ethical committee approved written consent form (Annexure I).

2.2. DNA isolation Protocol:

A. To the blood sample, added 4 volumes of reagent A in a polypropylene tube. Mixed gently till the solution becomes clear.

Reagent-A

Sucrose	109.54 gms	(320 mM)
1 M MgCl2	5.0 ml	(5 mM)
Triton X	10.0 ml	(1% Triton X 100)
1 M Tris-HCl (pH-8)	10.0 ml	(10mM)

Make up to 1000 ml with double distilled water (DDW). Autoclave.

B. Centrifuged at 2500rpm for 5 minutes to obtain a pellet, free from RBCs. The supernatant containing lysed RBCs is discarded carefully.

C. Disturbed the pellet thoroughly and added half the volume (as that of blood sample) of reagent B. Mixed thoroughly and gently by inverting for 3-4 minutes till the solution becomes viscous.

Reagent-B

1 M Tris-HCl (pH-8)	40 ml	(400mM)
0.5 M Na-EDTA	12 ml	(60mM)
1 M NaCl	15 ml	(150mM)

Make up to 95 ml with DDW, autoclave. Then add 5 ml 20% SDS (1%).

D. Add reagent C (1/4th volume of reagent B) and mix gently for 3-4 minutes.

Reagent-C

5 M Na-per chlorate (MERCK)- 100 gms.

Make up to 142 ml with DDW (do not autoclave)

E. Add equal volumes (as that of regents B+C) of Phenol + Chloroform + Isoamyl alcohol (25:24:1). Mix well and centrifuge at 2500 –3000 rpm for 7-8 minutes to separate into 3 layers viz. aqueous layer, protein layer and solvent layer.

F. Transfer the aqueous layer carefully into another centrifuge tube using a broad mouth tip (Care should be taken that the protein layer is not disturbed).

G. Add equal volumes of Chloroform to the supernatant and mix gently for a minute and centrifuge at 2500 rpm for 5 minutes.

H. Transfer the aqueous phase to a fresh tube.

I. Add 2 volumes of chilled absolute alcohol and mix gently to precipitate the DNA.

J. Spool out the DNA lump in a fresh Eppendorf tube and decant alcohol.

K. Wash the DNA twice with 70%alcohol and give a short spin to remove alcohol.

L. Dry the pellet properly and ensure that whole alcohol is dried off.

M. Dissolve the pellet in 50-100ul of TE.

TE-Buffer

1 M Tris-HCl 1.0 ml (10.0mM)

0.5 M EDTA 0.2 ml (1mM)

Make up to 100 ml with DDW

N. Incubate at 55^0 C for 45 minutes to enhance the dissolution and store the DNA samples at 4^0C

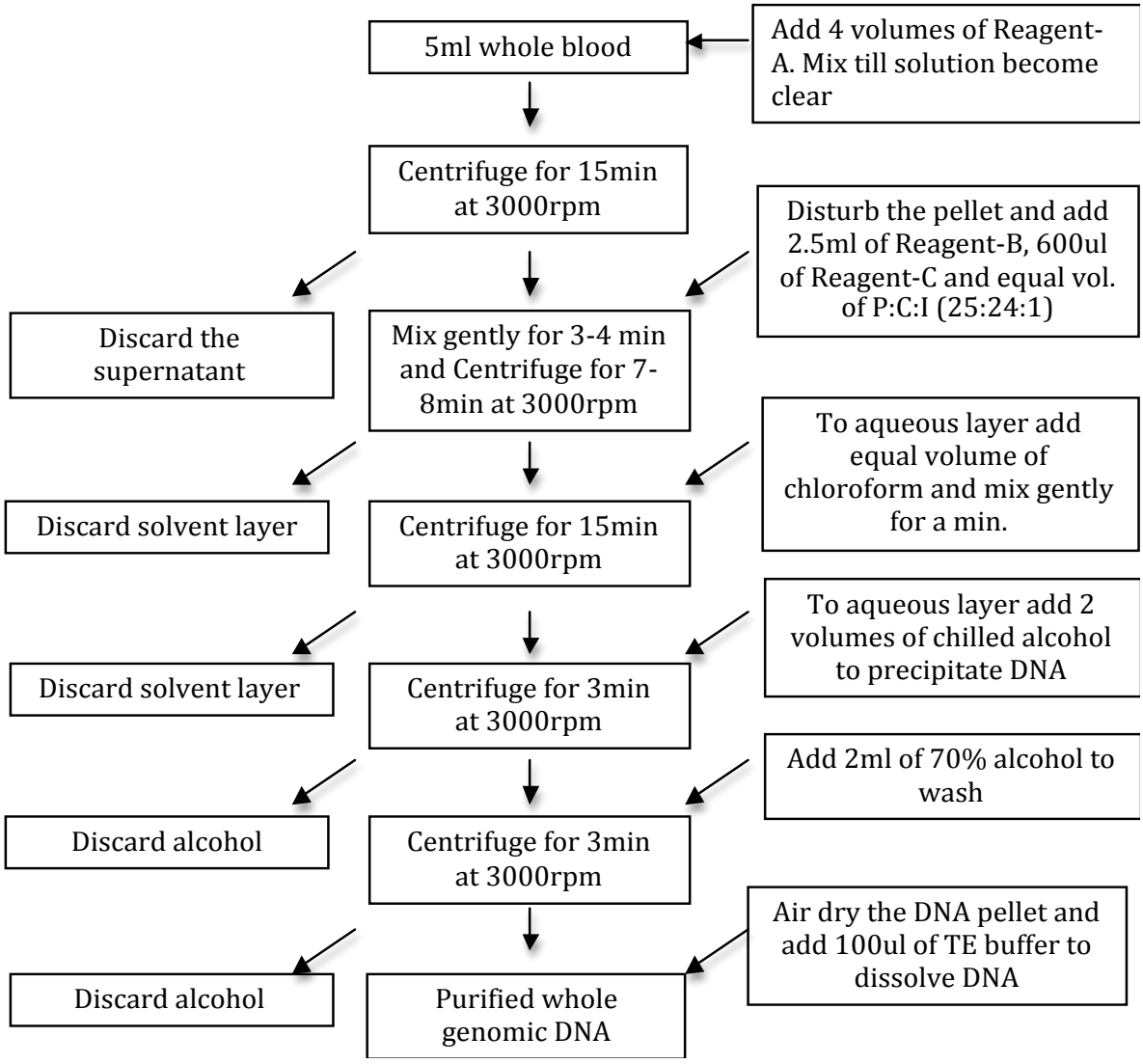
5ml whole blood
Add 4 volumes of Reagent-A. Mix till solution become clear
Centrifuge for 15min at 3000rpm
Disturb the pellet and add 2.5ml of Reagent-B, 600ul of Reagent-C and equal vol. of P:C:I (25:24:1)
Discard the supernatant
Mix gently for 3-4 min and Centrifuge for 7-8min at 3000rpm
To aqueous layer add equal volume of chloroform and mix gently for a min.
Discard solvent layer
Centrifuge for 15min at 3000rpm
To aqueous layer add 2 volumes of chilled alcohol to precipitate DNA
Discard solvent layer
Centrifuge for 3min at 3000rpm
Add 2ml of 70% alcohol to wash
Discard alcohol
Centrifuge for 3min at 3000rpm
Air dry the DNA pellet and add 100ul of TE buffer to dissolve DNA
Discard alcohol
Purified whole genomic DNA

Figure 3: Workflow of DNA isolation

2.3. DNA quantification:

Nucleic acids absorb light at a wavelength of 260 nm. For double-stranded DNA, an Optical Density (OD) of 1 at 260 nm correlates to a DNA concentration of 50 ng/µl, so DNA concentration can be easily calculated from OD measurements. These measurements were taken on a NanoDrop that requires only 1 µl of a sample for quantification.

2.4. Microarray

To gain insights, 50 samples of diagnosed MI cases and 50 samples of healthy unrelated control samples were processed for the microarray using Bead Chip Array, Global Screening Array-24 + v1.0, HTS, GSA + Multi-Disease with a capacity of ~ 6,40,000 markers for each sample. Significant variants were validated in a larger cohort of cases and controls.

2.5. Genotyping

2.5.1. Genotyping of microarray significant markers by Tetra Primer PCR

The microarray significant SNP's found in discovery panel (Microarray of 100 Samples i.e. 50 cases and 50 controls) were replicated for validation on larger cohort of MI cases and healthy controls. To validate the findings a total of 832 subjects were recruited in replication panel including 400 cases and 432 controls, cases include MI diagnosed vs unrelated healthy controls. The discovery panel yielded the microarray significant SNPs and replication panel was used to validate the significant found SNPs on larger cohort with an aim to see the role of significant association with disease outcome.

2.5.2. Genotyping of candidate gene markers by Tetra primer PCR and Sanger sequencing

From the review of literature survey interesting candiadte gene markers were selected and validated on 832 subjects which include 400 cases and 432 controls.

2.6. Electrophoresis Tetra Primer PCR was carried out on all the samples, the normal and the mutant primer products of same sample was loaded in different wells on 2% Agarose gel in TAE buffer with 100bp ladder and after the run is completed at 100v for 45 minutes the gel picture was captured using gel documentation system. The image was immediately numbered sample wise in PowerPoint for further genotyping.

2.7. Statistical Analysis

Bivariate and multivariate logistic regression was used to analyse data in dominant, recessive and additive models, age, sex and biochemical parameters were used as co-factors in analysis of data. Odds ratios were calculated at 95% confidence level in all the three models i.e. Dominant, Recessive and additive for all the cases and controls using SPSS software package (v16.0, SPSS Chicago) *P*- values >0.05 were considered as significant and role of genotyped SNP (Causative/Protective) was estimated based on odds ratio.

3. RESULTS

The present study analysed the genetic risk factors for myocardial infarction in a total of 856 samples, comprised of 417 clinically diagnosed MI patients (males 337 and females 80) and 439 healthy controls (males 264 and females 175) with no history of cardiac issues (Table 8). We looked at genetic risk factors for MI and how they correlated with lipid profile variables including blood total cholesterol, triglycerides, HDL, LDL, and VLDL.

Table 1: Sample characteristics

	Cases (n=417)	Controls (n=439)	P value
Age	58.32 (± 10.69)	58.5 (± 9.25)	0.7894
Gender (M/F)	337/80	264/175	0.9979
Total Cholesterol	199 (± 50.67)	191 (± 52.23)	0.0275*
Triglycerides	158.4(± 92.39)	139.2(± 62.43)	0.0005***
HDL	45.32 (± 13.74)	52.24 (± 35.89)	0.0003***
LDL	125.7 (± 43.47)	109.4 (± 48.91)	<0.0001****
VLDL	31.69 (± 18.53)	29.13 (± 13.06)	0.1104

3.1. GWAS in a sub-set of MI case and control samples using Illumina microarray platform

The new variants identified by GWAS were microarray-tested in the discovery panel and then validated in the replication panel to establish their significance. After genotype calling with Illumina's GenomeStudio software, 98 out of a total of 100 samples were considered successful (Figure: 4). PLINK software and the R statistical programming language was used to transfer the genotyping data from the samples with a call rate of 90–99% into a format suitable for further analysis.

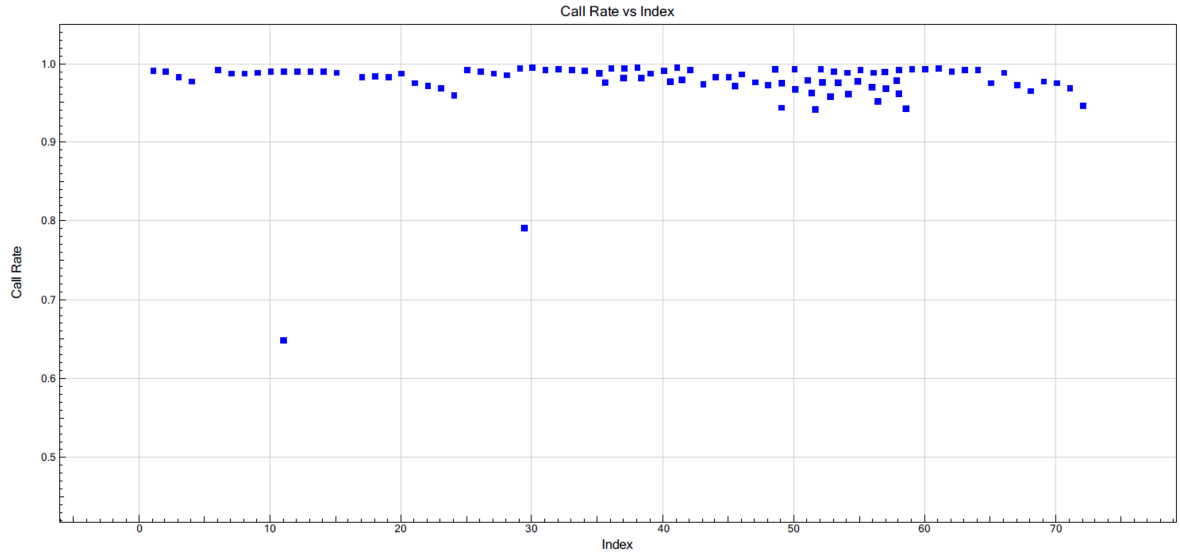

Figure 4: Call rate for 100 samples used in microarray

3.2. Identification of variants may be responsible for Mi from GWAS data

Following the techniques outlined in Clarke G et al., 2011, the resulting GWAS data was further analysed to determine the association status of markers. To check for potential systematic departures from exception due to residual population stratification, a QQ plot of association was created (Figure: 5) using the genomic inflation factor. A Manhattan plot was constructed to visualise genome-wide data and genetic association signals (Figure 6).

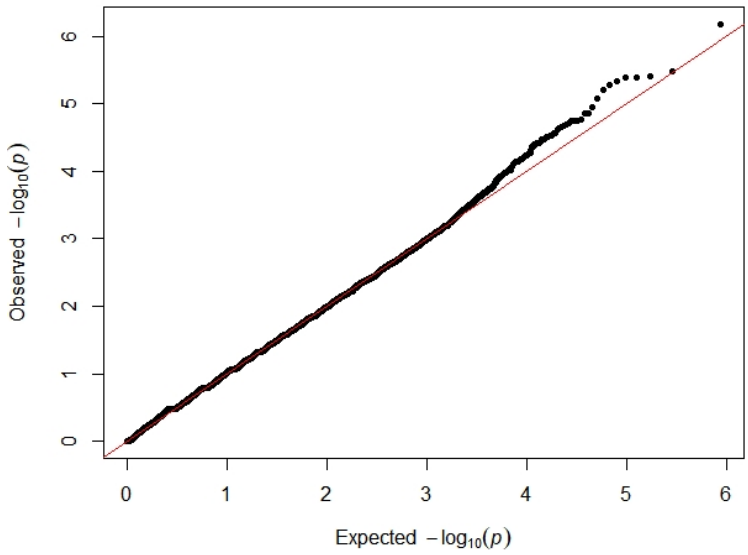

Figure 5: QQ plot showing association quality of samples

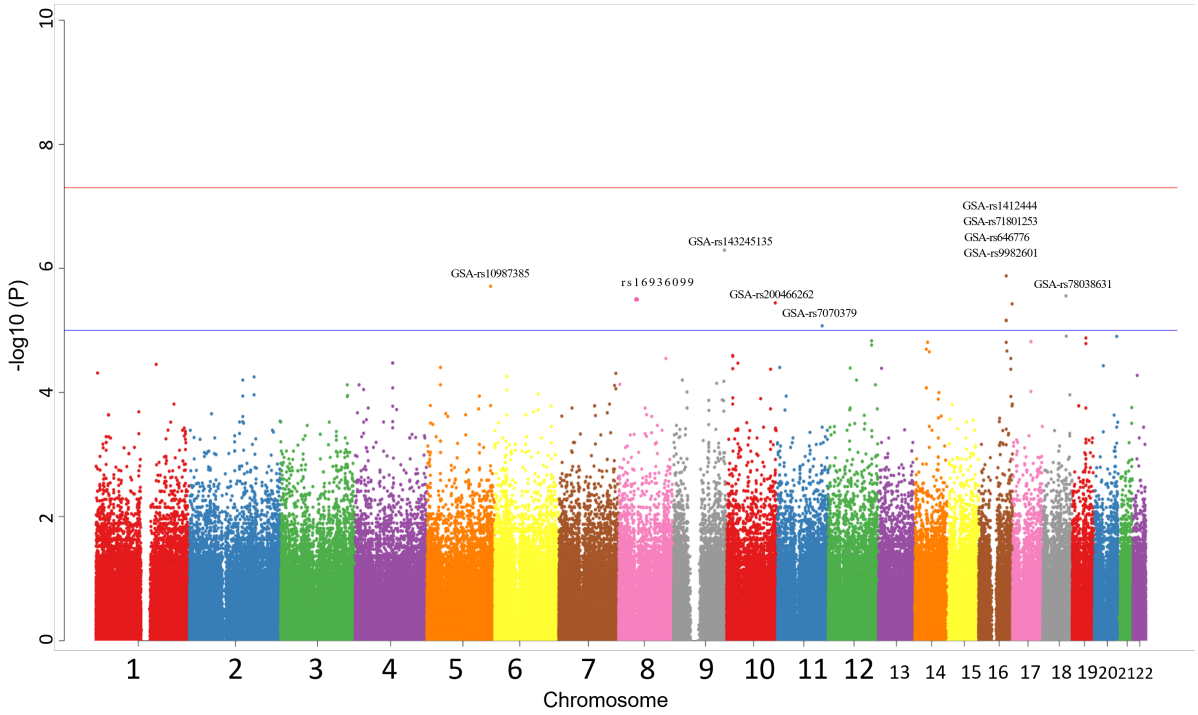

Figure 6: Manhattan plot showing association of genes with MI case and controls

As a result of the manhattan plot approach, nearly 10 markers have been identified as significan in the association study, in which rs16936099 of SULF1 gene and rs78038631 of KCNIP1 SNPs were reported earlier in association with myocardial infarction (Kumar et al., 2018 and Wang L et al., 2018).

3.3. Validation of MI associated variants from GWAS in a case-control cohort

3.3.1. SULF1 – rs16936099 validation

A significant SNP - rs16936099 of SULF1 gene was validated on 383 MI cases and 418 controls using the tetra primer ARMS-PCR method. Amplified products were run on a 2% agarose gel stained with EtBr for the amplicon separation and visualization on a gel documentation system. Genotypes were determined based on the amplicon size observed in the gel image. A 316bp band for the outer primers used for the targeted region amplicon, a 205bp band for the C allele specific amplicon and a 166bp band for the T allele specific amplicon were expected. The combination of 316bp and 205bp bands is considered the CC

genotype, the combination of 316bp and 166bp bands is considered the TT genotype, and if all three bands are observed the CT genotype (Figure: 7).

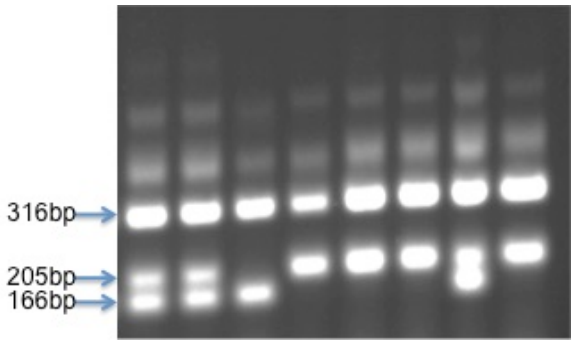

Figure 7: EtBr stained 2% agarose gel showing PCR amplified product of SULF1-rs16936099 marker allele specific amplicons.

The genotype information was used for further association studies in Grapghpad Prism and SPSS software. A dominant and recessive model association studies were also conducted along with genotype association. The analysis showed that the frequency of TT genotype is significantly higher in MI cases than in the controls (p = 0.0001****) (Table: 2). It is also found that the T allele is associated with an increased risk of MI (OR =1.587, CI 95% = 0.6274 – 20.002, p=0.0001****) (Table: 3).

Table 2: SULF1 rs16936099 genotype distribution of MI case-control

		Genotype frequency		**Control vs Case**	
Gene	**Genotype**	**Control (n=418)**	**Case (383)**	**Chi-square**	**P**
SULF1 rs16936099	CC	281 (67.2%)	230 (60.0%)	19.1	**<0.0001****
	CT	115 (27.5%)	98 (25.6%)		
	TT	22 (5.3%)	55 (14.4%)		
	CC Vs CT+TT (Dominant model)	137 (32.8%)	153 (40.0%)	4.452	**0.0349***
	CC+CT Vs TT (Recessive model)	396 (94.73%)	328 (85.63%)	19.0366	**0.000013*****

Table 3: SULF1 rs16936099 allelic distribution of MI case-control

	Allele frequency				**Control vs Case**		
Gene	**Control (n=836)**		**Case (n=766)**		**P**	**OR**	**95%CI**
SULF1 rs16936099	**C**	**T**	**C**	**T**	**0.0001*** **	1.587	0.6274 to 2.002
	677	159	558	208			

In a cases-control cohort, the association between lipid profile parameters and genotype frequencies was studied and no significant association was found (Figure 8, 9).

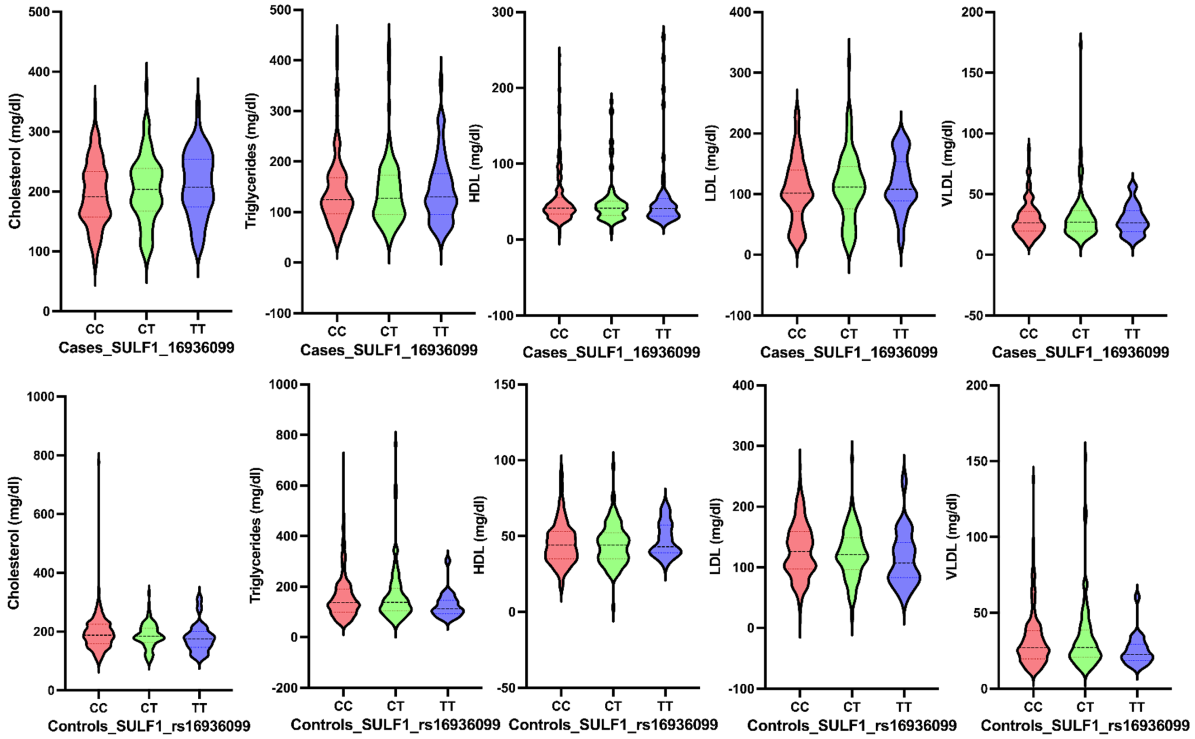

Figure 8: Violin plots showing no association of genotype Vs Case-controls Vs Lipoprotein parameters

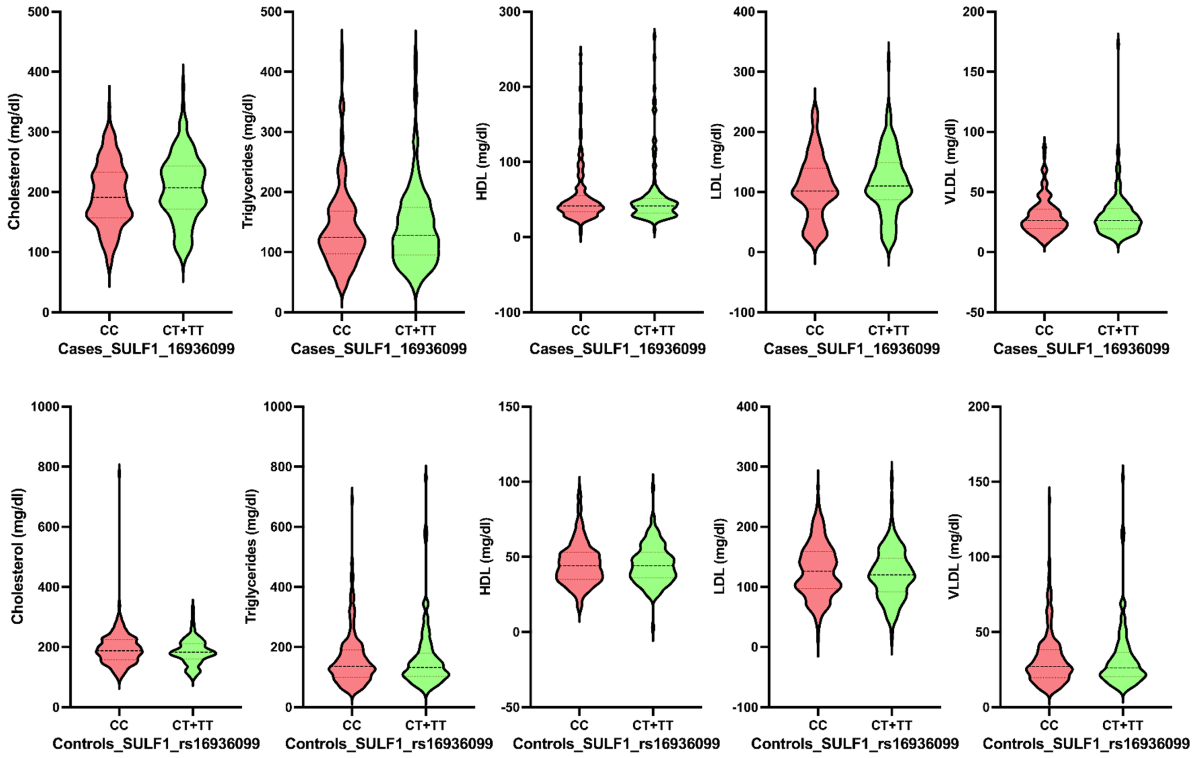

Figure 9: Violin plots showing no association of additive model genotype Vs Case-controls Vs Lipoprotein parameters.

3.4. Studying variations from candidate genes for their association with MI

To validate on MI cases and unrelated healthy controls, three SNPs were selected - rs6413504, rs1801706 and rs1050283, belonging to the LDLR, CETP and OLR1 genes respectively. Validation was done by Tetra-primer PCR and Sanger sequencing methods using custom designed primers, validated In-silico initially followed by proper standardization through gradient PCR.

3.4.1. LDLR – rs6413504 validation

LDLR – rs6413504 SNP was validated on 400 MI cases and 404 unrelated healthy controls using the tetra primer ARMS-PCR method. PCR products were documented in the gel documentation system after electrophoresis on a 2% agarose gel stained with EtBr. Genotypes were determined based on the amplicon size, A 432bp band for the outer primers used for the targeted region amplicon, a 273bp band for the G allele specific amplicon and a

217bp band for the A allele specific amplicon. The combination of 432bp and 273bp bands is considered the GG genotype, the combination of 432bp and 217bp bands is considered the AA genotype, and AG genotype is considered when all three bands are observed (Figure: 10).

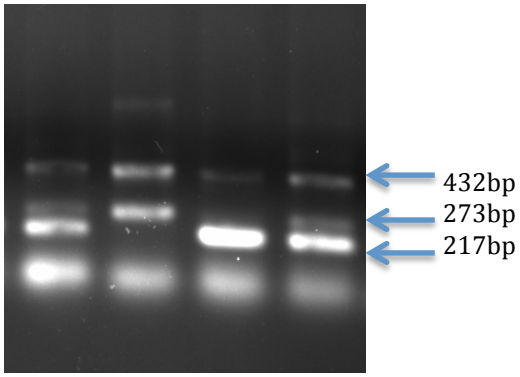

Figure 10: EtBr stained 2% agarose gel showing PCR amplified product of LDLR – rs6413504 marker allele specific amplicons.

The genotype data was analyzed using Grapghpad Prism software. Genotype association with the minor allele and major allele, dominant and recessive model association studies were conducte. The analysis showed that both genotype and allelic frequencies not shown any association with MI (Table: 4 & 5).

Table 4: LDLR rs6413504 genotype distribution of MI case-control

		Genotype frequency		**Control vs Case**	
Gene	**Genotype**	**Control (n=404)**	**Case (n=400)**	**Chi-square**	**P**
LDLR rs6413504	AA	97 (24.0%)	97 (24.3%)	1.575	0.4549
	AG	237 (58.7%)	248 (62.0%)		
	GG	70 (17.3%)	55 (13.7%)		
	AA Vs AG+GG (Dominant model)	307 (76.0%)	303 (75.7%)	0.006329	0.9366
	AA+AG Vs GG (Recessive model)	334 (82.67%)	345 (86.25%)	1.9584	0.1616

Table 5: LDLR rs6413504 allelic distribution of MI case-control

	Allele frequency				**Control vs Case**		
Gene	**Control (n=808)**		**Case (n=800)**		**P**	**OR**	**95%CI**
LDLR rs6413504	A	G	A	G	0.4424	0.926	0.7613 to 1.126
	431	377	442	358			

Lipid profile parameters were also checked for their association with the genotype data and a dominant model of genotype data and showed no significant association (Figure: 11 & 12)

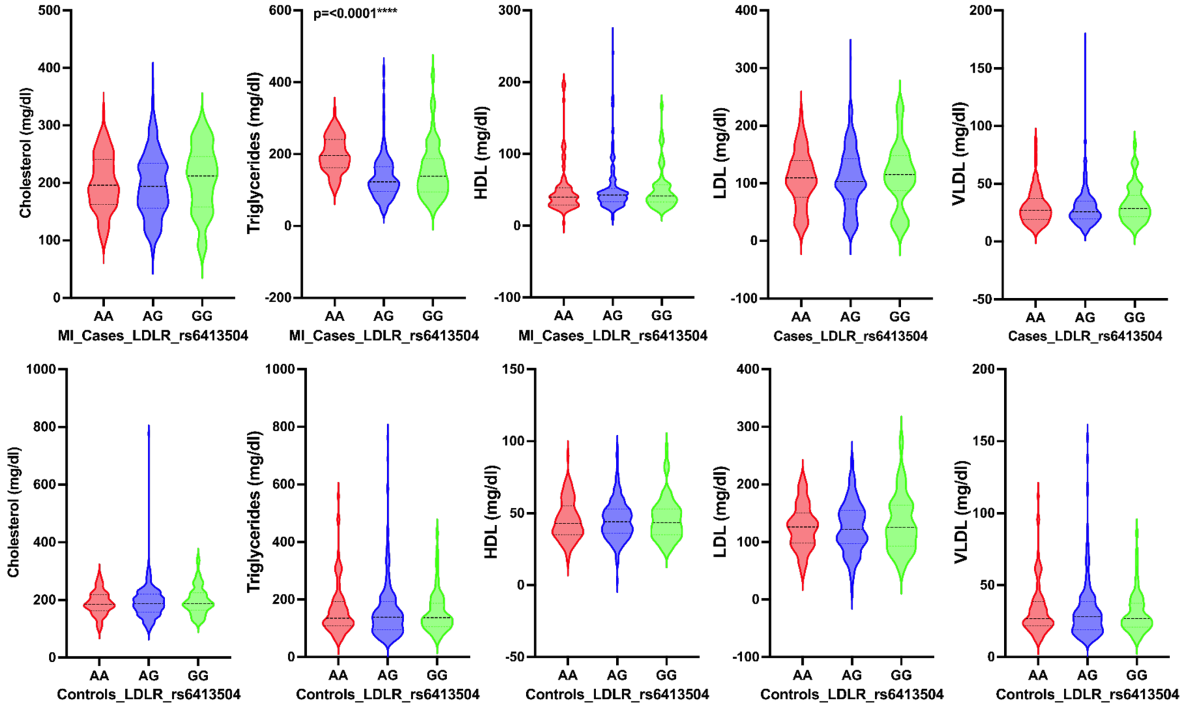

Figure 11: Violin plots showing no association of genotype Vs Case-controls Vs Lipoprotein parameters

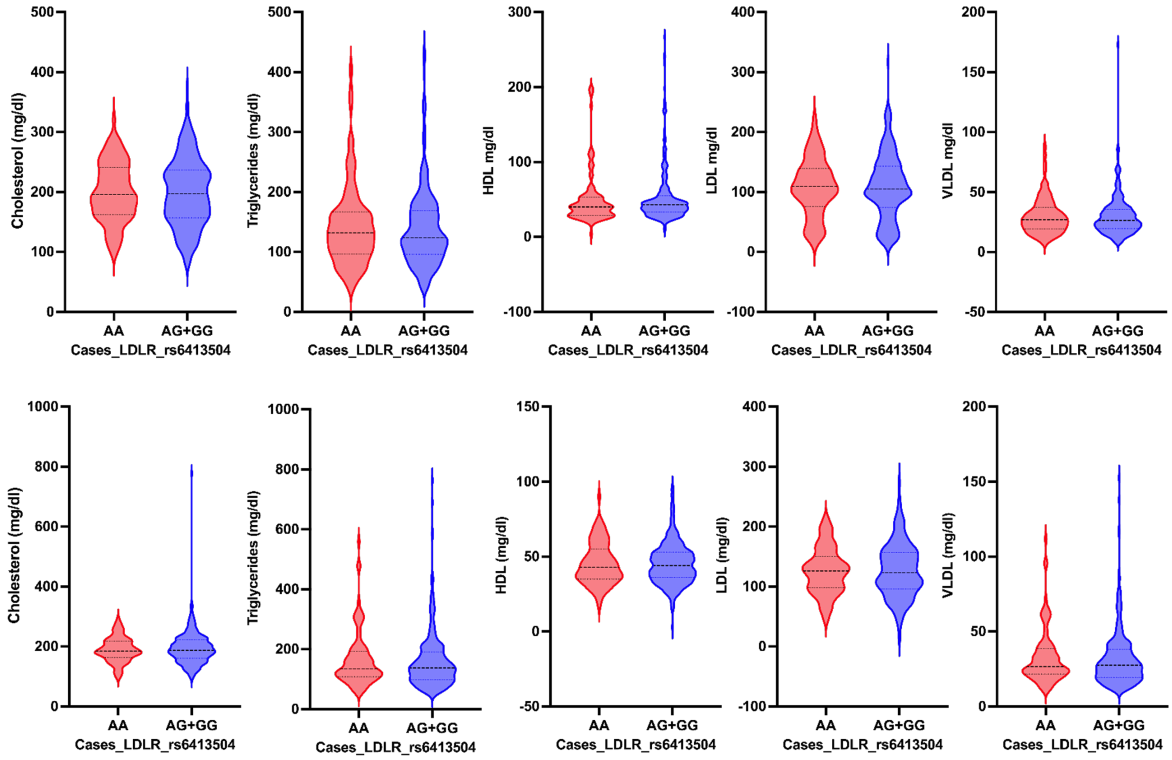

Figure 12: Violin plots showing no association of additive model genotype Vs Case-controls Vs Lipoprotein parameters

3.4.2. CETP – rs1801706 validation

The CETP – rs1801706 (c.84G>A) SNP was validated by Sanger's sequencing approach in 407 MI cases and 410 healthy controls. Sequencing data was analyzed for genotype variations by observing the electrophorograms (Figure: 13).

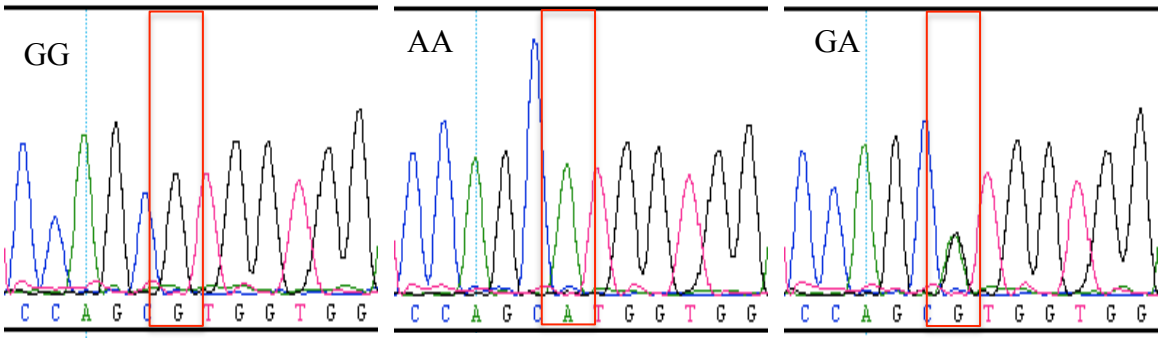

Figure 13: Sanger' sequencing of CETP-3'UTR G>A (rs1801706) PCR products representative electrophorograms of genotype CC—Homozygous normal, CT—Heterozygous, TT—Homozygous variant.

To understand the association status of the SNP with MI in case-control samples x^2 test was used. The analysis showed that the frequency of AA genotype is significantly higher in MI cases than in the controls (p = 0.01334**) (Table: 6). A allele showed the significant association with an increased risk of MI (OR =1.338, CI 95% = 1.042 – 1.711, p=0.0221*) (Table: 7).

Table 6: CETP rs1801706 genotype distribution of MI case-control

		Genotype frequency		**Control vs Case**	
Gene	**Genotype**	**Control (n=410)**	**Case (n=407)**	**Chi-square**	**P**
CETP rs1801706	GG	284 (69.3%)	260 (63.9%)	7.018	**0.0299***
	GA	116 (28.3%)	123 (30.2%)		
	AA	10 (2.4%)	24 (5.9%)		
	GG Vs GA+AA (Dominant model)	126 (30.7%)	147 (36.1%)	2.663	0.1027
	GG+GA Vs AA (Recessive model)	400 (97.56%)	383 (94.1%)	6.1229	**0.01334***

Table 7: CETP rs1801706 allelic distribution of MI case-control

	Allele frequency				**Control vs Case**		
Gene	**Control (n=820)**		**Case (n=814)**		**P**	**OR**	**95%CI**
CETP rs1801706	G	A	G	A	**0.0221***	1.338	1.042 to 1.711
	684	136	643	171			

The genotype data and dominant model of genotype data with lipid profile parameters showed no significant association (Figure: 14 & 15)

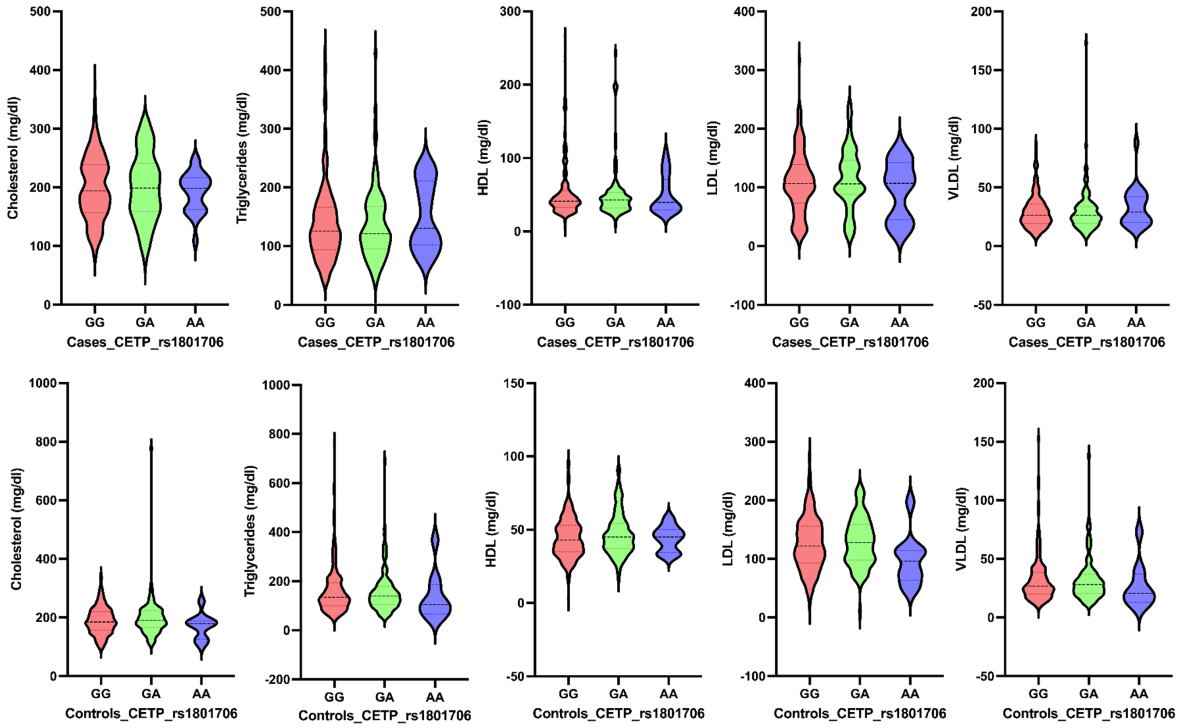

Figure 14: Violin plots showing no association of genotype Vs Case-controls Vs Lipoprotein parameters

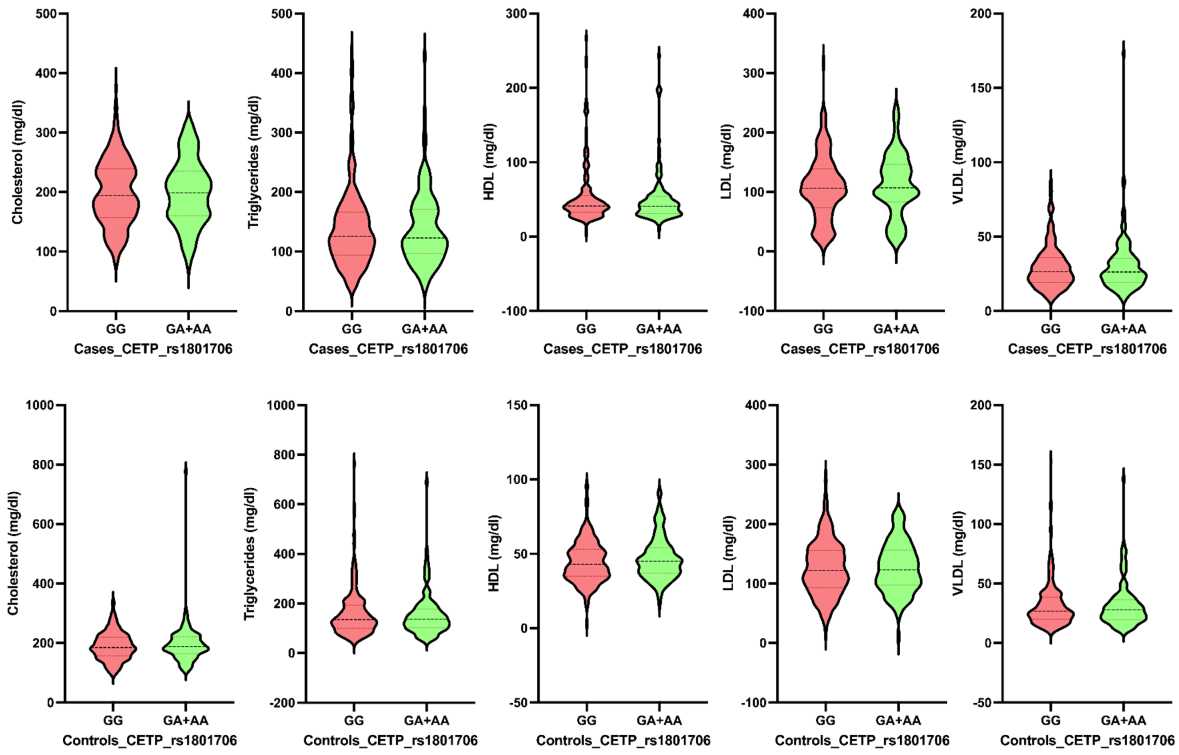

Figure 15: Violin plots showing no association of additive model genotype Vs Case-controls Vs Lipoprotein parameters

3.4.3. OLR1 – rs1050283 validation

OLR1 gene 3'UTR SNP rs1050283 was explored by Sanger's sequencing approach in 385 MI cases and 385 healthy controls. All samples were amplified for the targeted region comprising the SNP and checked on 2% agarose gel (Figure: 16).

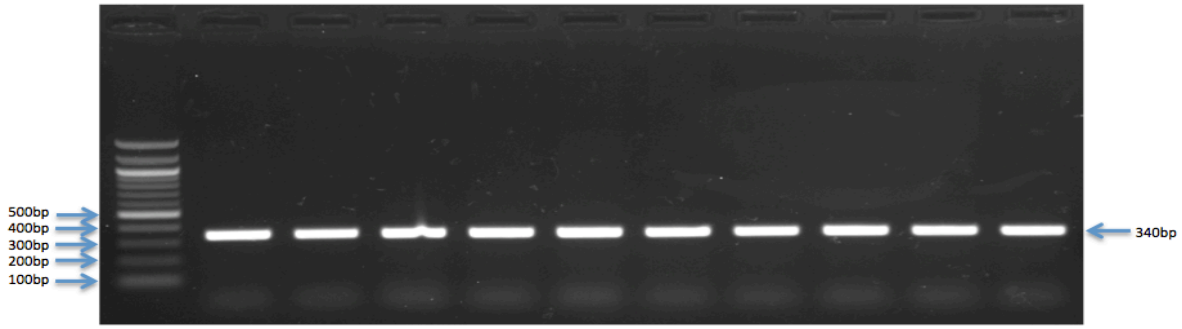

Figure 16: EtBr stained 2% agarose gel showing 340bp PCR amplified product of OLR1-3'UTR C>T (rs1050283).

Sequencing data was analyzed for genotype variations at SNP site in the targeted region of the gene by observing the electrophorograms (Figure: 17).

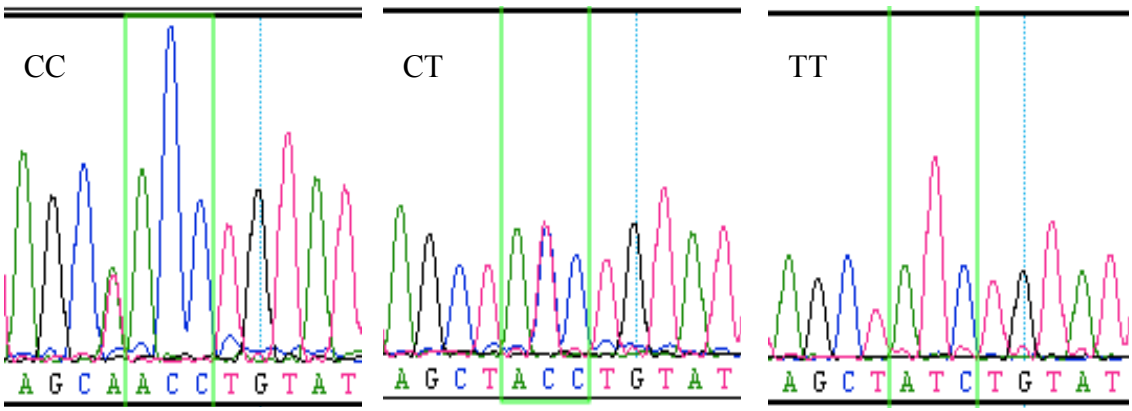

Figure 17: Sanger' sequencing products representative electrophorograms of genotype CC—Homozygous normal, CT—Heterozygous, TT—Homozygous variant.

The association status in MI case-control samples x2 test was used. The analysis showed that the frequency of AA genotype is significantly higher in MI cases than in the controls (p = 0.01334**) (Table: 8). A allele showed the significant association with an increased risk of MI (OR =1.338, CI 95% = 1.042 – 1.711, p=0.0221*) (Table: 9).

Table 8: OLR1 rs1050283 genotype distribution of MI case-control

		Genotype frequency		Control vs Case	
Gene	**Genotype**	**Control (n=385)**	**Case (n=385)**	**Chi-square**	**P**
OLR1 rs1050283	CC	243 (63.1)	223 (58.0)	6.756	**0.0341***
	CT	132 (34.3)	138 (35.8)		
	TT	10 (2.6)	24 (6.2)		
	CC Vs CT+TT (Dominant model)	142 (36.9)	162 (42.0)	2.174	0.1403
	CC+CT Vs TT (Recessive model)	375	361	6.031	**0.01405****

Table 9: OLR1 rs1050283 allelic distribution of MI case-control

	Allele frequency				Control vs Case		
Gene	**Control (n=770)**		**Case (n=776)**		**P**	**OR**	**95%CI**
OLR1 rs1050283	C	T	C	T	**0.0363***	1.295	1.015 to 1.645
	618	152	584	186			

Genotype data and dominant model genotype data with lipid profile parameters have not shown any association (Figure: 18 & 19)

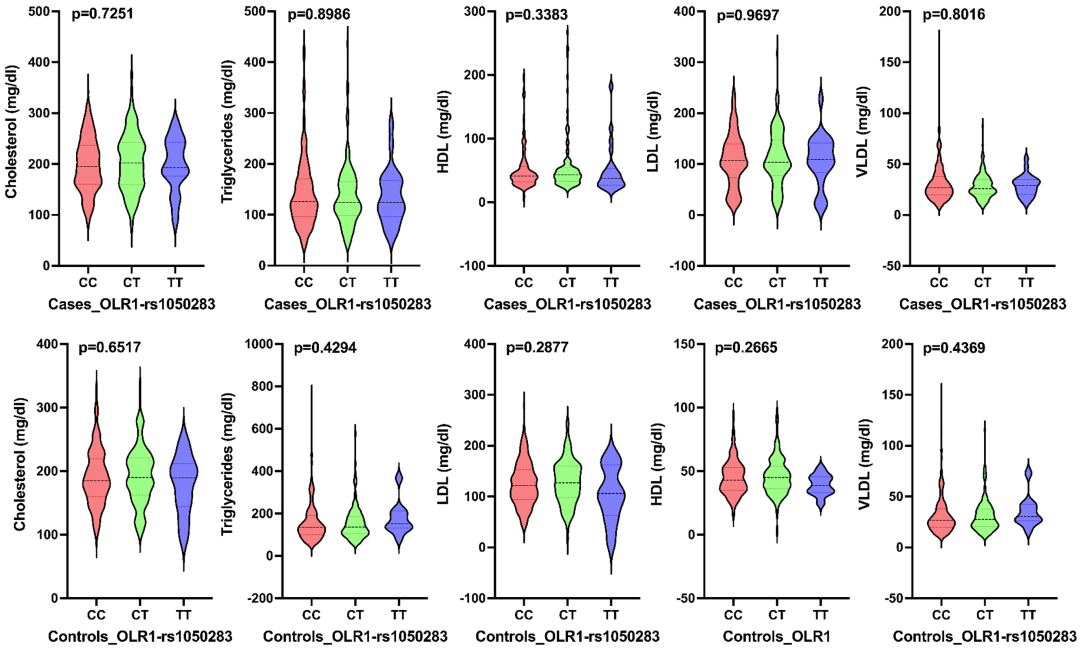

Figure 18: Violin plots showing no association of genotype Vs Case-controls Vs Lipoprotein parameters

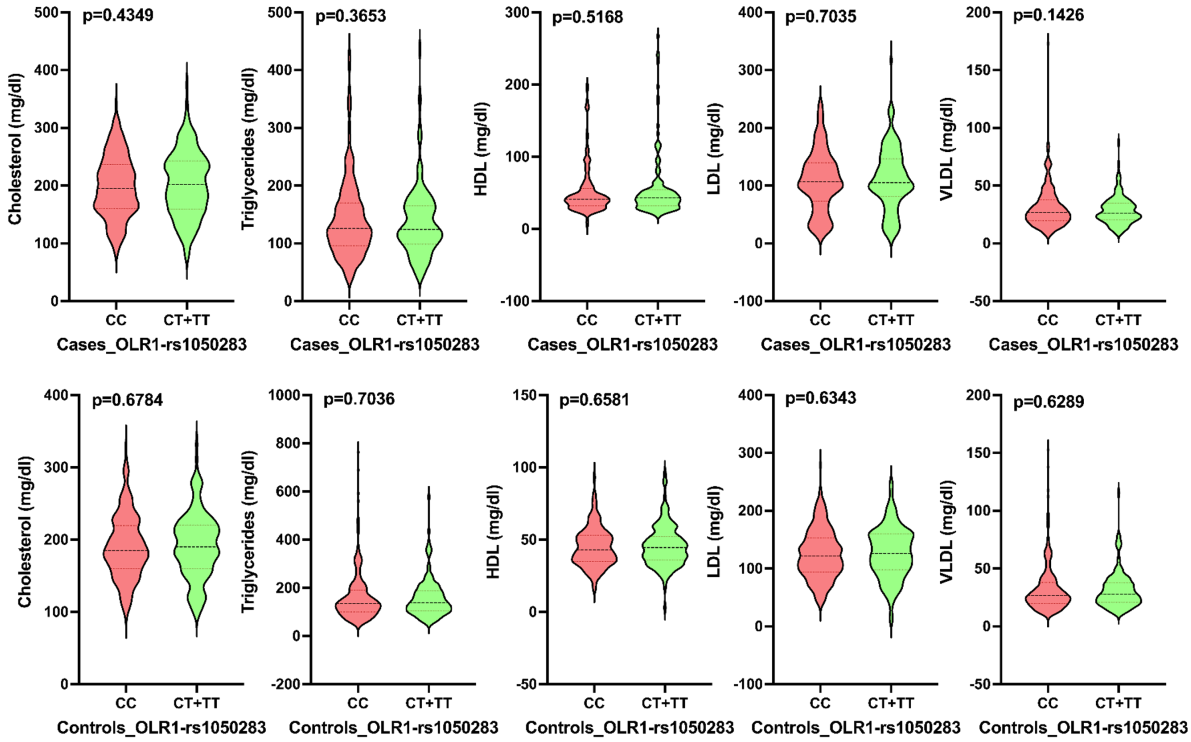

Figure 19: Violin plots showing no association of additive model genotype Vs Case-controls Vs Lipoprotein parameters

4. SUMMARY

A myocardial infarction (MI) results from a complicated interaction between an individual's genetic makeup and their exposure to several environmental circumstances, making it a polygenic, multi-factorial CVD. The precise causes of MI vary greatly depending on historical and regional factors. Millions of DNA sequence variants in the human genome have opened the door to research into the genetics of cardiovascular disease and other disorders across populations. Genome wide association studies (GWASs) have shown great promise in recent years in uncovering rare and common variations associated with numerous diseases, where the full genome is mapped and gene loci impacting susceptibility or resistance were later identified. Using a genome-wide method to improve the ability to identify new susceptibility genes might be a game-changer. Whole-genome analysis, in addition, will allow for the highest standard of inspection possible. Furthermore, the central mechanism in India, particularly in southern India, has not been studied.

In the current investigation, we used a GWAS and candidate gene strategy to investigate the link between a set of new and candidate gene markers and MI in a large sample cohort. There were 417 MI patients that were properly diagnosed and characterized, as well as 439 age-matched healthy controls. The phenol-chloroform extraction procedure was used to extract DNA from the blood samples of both cases and controls, and this DNA was then quantified. Until they were needed, the DNA samples were kept in a freezer at -200 degrees Celsius.

Bead Chip Array, Global Screening Array-24 + v1.0, HTS, GSA + Multi-Disease microarrays with a capacity of 6,40,000 markers were used to analyse 50 samples of identified MI cases and 50 samples of healthy, unrelated control samples for GWAS insights.

A larger cohort of MI cases and controls was used to duplicate and validate the microarray significant SNP's revealed in the discovery panel (Microarray of 100 Samples, or 50 cases and 50 controls). In addition, the significance of the rs16936099 of the SULF1 gene, T allele

(OR = 1.587, 95% CI = 0.6274-2.002, p 0.0001) and TT genotype with p 0.0001 was confirmed in a larger sample size of 801 participants, including 383 cases and 418 controls.

In this analysis, no significant connection between LDLR rs6413504 and MI was found. Those with MI had a genotype distribution of 260 GG, 123 AG, and 24 AA at the CETP gene's rs1801706/c.*84G>A, while controls had a distribution of 284 GG, 116 AG, and 10 AA at this same locus. Hardy-Weinberg equilibrium was detected for rs1801706/c.*84G>A in both cases and controls (2 p-value = 0.118 and 0.29, respectively). The rs1801706/c*84G>A genotype was found to be significantly associated with MI risk (p = 0.0299), where the AA and AG genotypes confer a higher relative risk of MI compared to the GG genotype. In a similar vein, the allelic distribution indicated 643 G and 171 A in cases, compared to 684 G and 136 A in controls. Significant differences in allelic distribution were found between patients and controls (2 p-value = 0.0221). Also, compared to the G allele, the A allele of rs1801706/c*84G>A increases the risk of MI.

The OLR1 gene was sequenced by Sanger and the 3'UTR C>T variant was analysed. There was little to no deviation from Hardy–Weinberg equilibrium in the frequency of individual alleles. T allele frequency: MI - 0.2416, control - 0.1974, p= 0.0363, odds ratio- 1.295, 95% CI: 1.015-1.645) and TT genotype distribution (x2 p 0.0341) were significantly different between the two groups.

There's more and more data suggesting that genetics have a significant impact in MI. Particularly intriguing are genetic factors because of the insight they provide into underlying biological circuits. Very few Indian studies have focused on the role that genes play in CVDs. This is the first study to our knowledge to investigate the genotype and allelic connection with MI in the South Indian population, as opposed to the North Indian population. Conclusion: People in the South Indian population who have mutations in several key genes involved in lipid metabolism are at increased risk for experiencing a heart attack.

5. REFERENCES

Erdmann, Jeanette, Patrick Linsel-Nitschke, and Heribert Schunkert. "Genetic causes of myocardial infarction: new insights from genome-wide association studies." *Deutsches Ärzteblatt International* 107, no. 40 (2010): 694.

Yamada, Yoshiji, Sahoko Ichihara, and Tamotsu Nishida. "Molecular genetics of myocardial infarction." *Genomic medicine* 2, no. 1-2 (2008): 7-22.

Ivanova, Anastasiya A., Vladimir N. Maksimov, Pavel S. Orlov, Dinara E. Ivanoshchuk, Sergei V. Savchenko, and Mikhail I. Voevoda. "Association of the genetic markers for myocardial infarction with sudden cardiac death." *Indian heart journal* 69 (2017): S8-S11.

Wang, Yongqin, Lefeng Wang, Xin Liu, Yongzhi Zhang, Liping Yu, Fan Zhang, Lisheng Liu, Jun Cai, Xinchun Yang, and Xingyu Wang. "Genetic variants associated with myocardial infarction and the risk factors in Chinese population." *PloS one* 9, no. 1 (2014): e86332.

Kontou, Panagiota, Athanasia Pavlopoulou, Georgia Braliou, Spyridoula Bogiatzi, Niki Dimou, Sripal Bangalore, and Pantelis Bagos. "Identification of gene expression profiles in myocardial infarction: a systematic review and meta-analysis." *BMC medical genomics* 11, no. 1 (2018): 1-11.

Schunkert, Heribert, Jeanette Erdmann, and Nilesh J. Samani. "Genetics of myocardial infarction: a progress report." *European heart journal* 31, no. 8 (2010): 918-925.

Shiffman, Dov, Stephen G. Ellis, Charles M. Rowland, Mary J. Malloy, May M. Luke, Olga A. Iakoubova, Clive R. Pullinger et al. "Identification of four gene variants associated with myocardial infarction." *The American Journal of Human Genetics* 77, no. 4 (2005): 596-605.

Matsuoka, Reiko, Shintaro Abe, Fumitaka Tokoro, Masazumi Arai, Toshiyuki Noda, Sachiro Watanabe, Hideki Horibe et al. "Association of six genetic variants with myocardial infarction." *International journal of molecular medicine* 35, no. 5 (2015): 1451-1459.

Winham, S. J., de Andrade, M. & Miller, V. M. Genetics of cardiovascular disease: Importance of sex and ethnicity. *Atherosclerosis.* (2015).

Gore, M. O., McGuire, D. K., Lingvay, I. & Rosenstock, J. Predicting cardiovascular risk in type 2 diabetes: the heterogeneity challenges. *Curr Cardiol Rep.* 17, 607 (2015).

Manchanda, S. C. & Madan, K. Yoga and meditation in cardiovascular disease. *Clin Res Cardiol.* 103, 675-680 (2014).

Tiago Justo, Antonie Martiniuc & Gurtej K. Dhoot. Modulation of cell signalling and sulfation in cardiovascular development and disease. *Sci. Reports.* 11.22424 (2021).

Holst, C. R. *et al.* Secreted sulfatases Sulf1 and Sulf2 have overlapping yet essential roles in mouse neonatal survival. *PLoS ONE* **2**(6), e575 (2007).

Ratzka, A. *et al.* Redundant function of the heparan sulfate 6-O-endosulfatases Sulf1 and Sulf2 during skeletal development. *Dev. Dyn.* **237**(2), 339–353 (2008).

Li Z, Zhao T, Tan X, Lei S, Huang L, Yang L. Polymorphisms in PCSK9, LDLR, BCMO1, SLC12A3, and KCNJ1 are Associated with Serum Lipid Profile in Chinese Han Population. Int J Environ Res Public Health. 2019 Sep 2;16(17):3207.

Tsompanidi, E.M., Brinkmeier, M.S., Fotiadou, E.H., Giakoumi, S.M. & Kypreos, K.E. HDL biogenesis and functions: role of HDL quality and quantity in atherosclerosis. *Atherosclerosis.* **208**, 3-9 (2010).

Drayna, D., Jarnagin, A.S., McLean, J., Henzel, W., Kohr, W., Fielding, C. et al. Cloning and sequencing of human cholesteryl ester transfer protein cDNA. *Nature.* **327**, 632-634 (1987).

Agellon, L.B., Quinet, E.M., Gillette, T.G., Drayna, D.T., Brown, M.L. & Tall, A.R. Organization of the human cholesteryl ester transfer protein gene. *Biochemistry.* **29**, 1372-1376 (1990).

Callen, D., Hildebrand, C. & Reeders, S. Report of the Second International Workshop on Human Chromosome 16 Mapping. *Cytogenetics and cell genetics.* **60**, 158 (1992).

Jiang, X.C., Moulin, P., Quinet, E., Goldberg, I.J., Yacoub, L.K., Agellon, L.B. et al. Mammalian adipose tissue and muscle are major sources of lipid transfer protein mRNA. *Journal of Biological Chemistry.* **266**, 4631-4639 (1991).

Inazu, A., Quinet, E.M., Wang, S., Brown, M.L., Stevenson, S., Barr, M.L. et al. Alternative splicing of the mRNA encoding the human cholesteryl ester transfer protein. *Biochemistry.* **31**, 2352-2358 (1992).

Brown, M.L., Inazu, A., Hesler, C.B., Agellon, L.B., Mann, C., Whitlock, M.E. et al. Molecular basis of lipid transfer protein deficiency in a family with increased high-density lipoproteins. *Nature.* **342**, 448-451 (1989).

Inazu, A., Jiang, X.-C., Haraki, T., Yagi, K., Kamon, N., Koizumi, J. et al. Genetic cholesteryl ester transfer protein deficiency caused by two prevalent mutations as a major determinant of increased levels of high density lipoprotein cholesterol. *Journal of Clinical Investigation.* **94**, 1872 (1994).

Kulsoom, B. & Hasnain, S.N. Association of serum C-reactive protein and LDL: HDL with myocardial infarction. *Journal-Pakistan Medical Association.* **56**, 318 (2006).

Research Scholar — Research Supervisor

Nagarjuna Pasupuleti — Prof. MS Mustak